正確使用統計圖表呈現處理間比較 [1]

呂秀英[2,3]

摘　要

呂秀英。2011。正確使用統計圖表呈現處理間比較。台灣農業研究 60:61–71。

精確表達資訊是統計圖表製作的首要原則。農業試驗研究最常見的資料分析，為利用變方分析 (analysis of variance, ANOVA) 及多重比較 (multiple comparisons) 來探討多組處理間的差異性。多重比較包含很多種方法，其中 Fisher 的最小顯著差異性測驗 (Fisher's protected least significant difference test, LSD test) 是農業學術期刊中最普遍接受的多重比較方法，主要用來進行所有處理平均值間的兩兩比較是否存有差異。這些統計分析的結果，若沒有在圖或表中予以適當呈現，會令人不易閱讀甚至產生誤解，進而影響研究報告發表的品質。因此，本文以實例說明如何用適當的圖表來有效呈現出 ANOVA 及 LSD 之分析結果，並給予精確的文字解讀，期能有助於農業科技研究成果發表之參考。

關鍵詞：變方分析、多重比較、成對處理平均值間比較、最小顯著差異性測驗。

農業試驗研究之目的，通常是在探討不同處理 (treatments) 間是否存在差異。變方分析 (analysis of variance, ANOVA) 及多重比較 (multiple comparisons) 是用來探討多組處理平均值間差異性的最常見統計方法之一。事先經適當試驗設計所獲得的研究資料，欲比較各處理平均值間之差異時，首先必須進行 ANOVA，一旦 ANOVA 表中處理效應的 F 測驗值顯著，表示處理平均值間存在顯著差異，但此時仍無法得知何種處理平均值顯著較大或較小；因此，還必須再進行多重比較，以探討兩兩處理間的個別差異性。多重比較是一群統計方法且迄今已發展出至少 20 多種分析方法 (Hsu 1996)，其中以 Fisher 的最小顯著差異性測驗 (Fisher's protected least significant difference test, LSD test) 最廣受生物農醫學領域所使用 (Lu 1990; Chiang 2006)。在文獻中常發現研究人員選用了正確的統計方法進行資料分析後，卻未能將分析結果適當的予以表達，除報告品質大打折扣外，也會造成讀者混淆甚至產生誤解；尤其統計圖表除數字外，也常用到一些符號，例如代表顯著性的星號 (*、**) 或英文字母 (a、b、…) 等，這些符號在圖表中必須適當標示並以清楚的文字註解，才能有效呈現出數字的含義。由於絕大多數學術期刊 (如本期刊) 均要求統計圖表的標題和內容必須以英文方式表達，這對於國內很多不甚精通統計學的農業研究人員而言，更是無所適從，於是針對圖表內符號意義的英文語法，只好按自身的文學造詣各自發揮，因而影響研究報告的品質。因此本文針對

1. 行政院農業委員會農業試驗所研究報告第 2478 號。接受日期：100 年 3 月 23 日。
2. 本所研究員兼副所長。台灣 台中市。
3. 通訊作者，電子郵件：iying@tari.gov.tw；傳真機：(04)23302801。

農業試驗資料進行 ANOVA 及多重比較後，如何正確運用圖表來有效呈現分析結果，並據以精確的文字解讀，以實例說明提供科研人員參考。

以單因子 CRD 試驗為例說明變方分析 (ANOVA)

ANOVA 的結果，其最終是以一個表格來呈現。無論以任一種試驗設計，如完全逢機設計 (complete randomized design, CRD)、逢機完全區集設計 (randomized complete block design, RCBD)、裂區設計 (split plot design, SPD) 等之 ANOVA 表都擁有相同的欄位：變因 (source)、自由度 (degrees of freedom, DF)、平方和 (sum squares, SS)、均方 (mean square, MS)、F 測驗值以及 F 之機率值 (Pr > F)，只是不同試驗設計之 ANOVA 所含的變因不盡相同而已 (Steel *et al*. 1997)。以單因子試驗之 CRD 的分析結果為例說明，假設針對某性狀進行 ANOVA 後，應將統計軟體輸出結果整理如表 1，表中可直接呈現出 F 之機率值 (Pr > F)，從表中的 Pr > F 值小於 0.01，可知該範例的處理效應 (treatment) 極顯著存在，表示該性狀在不同處理間的平均值差異已達 1%顯著水準 (significance level)。此外，也可以將上述 ANOVA 表再行簡化，省略 Pr > F 欄位而直接在 F 值旁，以星號 (*或**) 來表達顯著性結果，如表 2 所示，但此務必在表下加註星號的意義。統計上習慣以*及**各代表 Pr > F 小於 5%及 1%顯著水準，且各以顯著及極顯著來作文字上的解讀。倘若 ANOVA 表中同時出現*及**結果，則表下之英文註解應寫成：*,**Significant at 5% and 1% levels, respectively.

農業領域中的試驗往往會同時調查很多性狀，並就各性狀比較其在處理間之差異性。同一試驗多個性狀必須分別進行 ANOVA，因此統計軟體輸出結果就會產生多個 ANOVA 分析表，但在報告發表時為求簡潔並節省版面篇幅，不可能將每一個性狀皆以表 1 或表 2 逐一列出於報告中，而通常會將這些性狀的 ANOVA 分析結果整併成一個表格。因此，針對多種性狀之場合的 ANOVA 表，則可省略 SS 及和 F 值兩個欄位以及整行的總變因 (Total)，而只保留變因、DF 及 MS 欄，並將顯著性星號直接標示在 MS 值上，如以單因子試驗之 CRD 為例之表 3 所示。請注意，由於表中已略去 F 測驗值且直接將顯著星號標記於各 MS 值

表 1. 以 F 之機率值呈現顯著性程度的單因子試驗之 CRD 變方分析

Table 1. The ANOVA result in a single factor CRD experiment using probability of F value for the degree of significance

Source	DF	SS	MS	F	Pr > F
Treatment	3	11.7	3.90	9.75	0.00154
Error	12	4.8	0.40		
Total	15	16.5			

表 2. 以星號呈現顯著性結果的單因子試驗之 CRD 變方分析

Table 2. The ANOVA result in a single factor CRD experiment using asterisk(s) as level of significance

Source	DF	SS	MS	F
Treatment	3	11.7	3.90	9.75**
Error	12	4.8	0.40	
Total	15	16.5		

**Significant at 1% level.

表 3. 單因子試驗多種性狀之 CRD 變方分析

Table 3. The ANOVA results of different traits in a single factor CRD experiment

		Mean square			
Source	DF	Trait 1	Trait 2	……	Trait n
Treatment	3	3.90**	4.81*	……	300.2
Error	12	0.40	1.25	……	120.4

*,**Significant at 5% and 1% levels, respectively.

上，因此表下關於星號之意義的註解敘述中，宜特別強調顯著性的判斷是依據 F 值測驗。根據表 3 範例所呈現的結果可解讀為：Trait 1 在處理間呈極顯著差異、Trait 2 在處理間呈顯著差異、….、Trait n 在處理間之差異未達 5%顯著水準等。但從 ANOVA 表中並無法得知該等性狀到底是何種處理之平均值顯著較大或較小，必須再進行後續的多重比較方可獲悉。至於 ANOVA 表中各性狀的 MS 值應取至小數點後幾位數字，原則上以能充分顯現出處理項 (Treatment) 和機差項 (Error) 兩者之 MS 的差異性，但不得因四捨五入而改變兩者相除後所得 F 值之顯著性，且不得使機差項 MS 變為零。一般多將 MS 取至小數點後兩位數字，如表 3 範例中的 Trait 1 和 Trait 2 兩性狀之 MS；但當 MS 值很大時，為使表格版面看起來較簡潔，可在不違背上述原則下四捨五入再進位，如表 3 範例中的 Trait n 之 MS 取至小數點後一位數字，或甚至成為整數。ANOVA 表中若含有多種性狀，則不同性狀之 MS 值的呈現位數未必要一致，但同一性狀內各變因之 MS 值的位數務必相同。

在正式發表時為了節省期刊印刷篇幅，可省略單因子試驗之 ANOVA，直接以多重比較結果來呈現。但倘若是複因子試驗 (factorial experiment)，則絕對不可省略 ANOVA 結果！複因子試驗進行 ANOVA，最重要的是要從中先行判斷因子間交感效應 (interaction effect between the factors) 是否存在，因子之間交感效應存在與否，決定了後續的多重比較之分析方式也有所不同。例如 A、B 兩因子試驗，就 A 而言，B 為交感因子，反之就 B 而言，A 為交感因子，若交感效應不顯著，表示某因子對性狀的影響狀況不會隨著另一交感因子而異，故可就各因子分別執行多重比較；若交感效應顯著，表示某因子對性狀的影響狀況會隨著另一交感因子變級 (level) 的變化而有不同表現，因此可視情況選擇下列任一分析方式：(1) 將所有 A 因子與 B 因子組合之間互作多重比較，以從中篩選出最佳組合；(2) 固定某因子的變級，對另一因子進行多重比較，以釐清在固定某因子下另一因子變級間對性狀的影響差異，並藉此選出最佳組合；(3) 若兩因子皆為數量因子，可利用迴歸技術 (regression techniques) 尋找性狀觀測值對因子的關係式。一旦報告中沒有將複因子試驗之 ANOVA 列出來，就無法讓讀者了解交感效應是否存在，從而無法確認作者進行多重比較分析的適當性。

以 LSD 測驗為例說明多重比較

已知至少有 20 多種多重比較分析方法，分別適用於成對處理間比較 (pairwise comparisons of treatments)、非成對處理間比較 (unpairwise comparisons of treatments) 與控制組之比較 (comparisons with a control) 等 (Hsu 1996)。農業試驗研究上最常進行成對處理間比較，即所有處理平均值間的兩兩比較，而在多重比較法中最廣為所知的成對處理間比較法有 LSD 測驗、鄧肯氏新多重變域測驗 (Duncan's new multiple-range test, DMRT) 等，但由於 DMRT 現今已被公認並不適用，其原先認定的應用場合，建議最好改用無計畫設計比較的 LSD 測驗或有計畫設計之定量處理比較的迴歸技術 (Chew 1980; Carmer & Walker 1985; Lu 1990; Chiang 2006)；此相關議題，筆者率先在國內「科學農業」期刊探究此一問題 (Lu 1990a)，繼後國立中興大學蔣國司教授在「植物保護學會會刊」亦有詳細論述 (Chiang 2006)。因此強烈建議農業試驗在成對處理間比較之場合，使用 LSD 測驗是最安全的作法，此也是國際上很多 SCI 學術期刊所能接受的唯一方法。國內外許多學術期刊在其稿約中也已規範，不再接受多重比較法中之 DMRT 的使用。

研究人員如果認定過去類似試驗的成果報告都是採用 DMRT 進行處理間比較，所以仍沿襲同樣的方法，這是很不恰當的。

以多種性狀的 LSD 分析結果為例，整理成如表 4 所示，並掌握以下幾個重要原則：

(1) 各欄位的名稱字首要用大寫英文字母，例如 Treatment、Trait 1、Trait 2 等。

(2) 各性狀名稱旁應註明測量單位，例如 cm、%、g 等，但 number 除外。

(3) 除平均值（$\bar{x}$）外，宜同時列出分散度介量 SE 或 SD，如 $\bar{x} \pm$ SE 或 $\bar{x} \pm$ SD。SE 為平均值標準差 (standard error of mean)，SD 為標準差 (standard deviation)，兩者的公式和意義並不相同，一般在進行處理平均值比較時，多採用 SE (Streiner 1996; Lu *et al*. 2005)。

(4) 各性狀之平均值的有效位數必須配合分散度介量 (如 SE 或 SD) 的精準度並維持兩者一致，例如考慮 SE 或 SD 之精準度而將其有效位數取至小數點後一位數字時，則平均值的有效位數也要取至小數點後一位，因此寫成 12.1 ± 1.3 是適當的，切勿寫成 12.1 ± 1.32；若 SE 或 SD 的有效位數是取至小數點後兩位，則應寫成 12.10 ± 1.32 以顯示平均值與分散度介量之位數相同。而考量各性狀之測量單位及平均值數字大小的差異，不同性狀平均值間的有效位數未必要求一致，原則上以能充分顯現出 SE 或 SD 在處理之間的差異性即可，例如表 4 範例中的性狀 Trait n，將 SE 取至整數已能充分表達其在處理間之差異，故未必要取至小數點後一位；反之，若將 Trait 1 的 SE 取至整數，則可發現該性狀所有處理的 SE 值都因四捨五入進位而成為 1，此時便無法顯現出各處理 SE 的差異，必須至少取至小數點後一位才能突顯出處理間 SE 的差異，但若將其取至小數點後兩位，卻可能因為小數點後的數字太多，反而變得繁瑣讓整個表格變得不易閱讀且未必適當。換言之，統計表內數字位數之呈現，必須同時考量在闡明差異性需求與維持版面精簡可讀之間取得平衡。另外，表中同一欄內所有數字應一律以小數點對齊，這是一個好的統計表之基本要求。

(5) 成對處理間之差異顯著性，習慣上多以平均值最大者給予 a，依次給 b, c…等之小寫英文字母來區分。少數期刊並未對此有嚴格統一規定，故偶爾含有以平均值最小之處理給予 a 的報告，對此筆者強烈建議不要採納這種作法，以確保研究報告之品質。此外，顯著性字母與數字之間宜用一個空格區分開來，但無須將字母使用上標，以免因字體過小而導致閱讀不易。

(6) 表下應以文字加註 SE (或 SD)、重複次數 (n) 及顯著性字母的意義，並說明使用何種多重比較方法。圖表中有關顯著性字母的意義註解，可採用以下不同寫法：

> Means within each column followed by the different letter(s) are significantly different at $P < 0.05$ by Fisher's protected LSD test.
>
> Means within each column followed by the same letter(s) are not significantly different at $P < 0.05$ by Fisher's protected LSD test.
>
> Means within each column followed by the different letter(s) are significantly different at 5% level by Fisher's protected LSD test.
>
> Means within each column followed by the same letter(s) are not significantly different at 5% level by Fisher's protected LSD test.

應當注意若有多個性狀並列之場合，處理間是逐欄比較時，表下英文註解要寫清楚是 within each column，否則讀者會誤認是逐行比較；倘若處理間是逐行比較時，則英文註解要改寫成 within each row；但逐行比較時，由於各處理的

顯著性字母之間被平均值隔開，不易讓讀者馬上能看出各處理間的差異顯著性，因此通常以逐欄比較為常見表達方式，其同一性狀各處理的顯著性字母列在同一欄內，可有利於比較；但針對單一表格內同時有逐欄與逐行兩種不同比較方式的場合，為避免混淆不清，逐欄及逐行比較可分別採用小寫及大寫的英文字母來作為顯著性符號，如表 5 及表 6 範例所示，至於小寫、大寫字母何者在前，並沒有固定用法，只要讓表格內的顯著性結果可以一目了然即可。再者，圖表註解行內無論是強調相同字母或不同字母 (不顯著或顯著)，使用如表 4 及表 5 之 5% level (即 0.05 probability level) 或如表 6 之 $P < 0.05$ 的顯著性基準皆可，但切勿寫成 $P \leqq 0.05$ 或 $P > 0.05$。在有些報告中常見將 Fisher's protected LSD test 簡單寫成 LSD test，這並不會構成太大的問題，不過一旦報告中省略未列 ANOVA 而只呈現 LSD 結果時，建議在名稱上最好還是維持有 protected 這個字。這是因為 Fisher 的早期作法，是可以未進行 ANOVA 之 F 測驗就利用 LSD 比較平均值，但此時設定相同顯著性水準下兩者統計上的犯錯率 (error rate) 會不一致，而有可能產生 F 與 LSD 測驗的結果會有差異，即出現 F 不顯著但 LSD 顯著的現象，因此在未先執行 ANOVA 就進行 LSD 測驗時必須修正顯著水準，此所謂 unprotected LSD；若在 F 檢定後再以 LSD 比較，兩者犯錯率接近，因此 protected LSD 表示是有先經過 ANOVA 的 F 測驗後再就其顯著與否，來決定多重比較是否值得進行 (Hsu 1996; Steel *et al.* 1997)。當 ANOVA 之 F 測驗不顯著時，表示所有處理間未達顯著差異，因此後續的 LSD 結果自然可省略不必執行，而直接在平均值比較表內將各處理平均值給予相同的字母 a 即可。

根據表 4 範例的結果可解讀為：經 LSD 測驗在 5%顯著水準下，性狀 Trait 1 以處理 T1 和 T2 之平均值皆顯著高於 T3 和 T4，T3 和 T4 間則差異未顯著；性狀 Trait 2 在四個處理間未有顯著差異；性狀 Trait n 以處理 T1 和 T2 之平均值皆顯著高於 T4，處理 T1 亦顯著高於 T3，但

表 4. 多種性狀在不同處理間之比較

Table 4. Comparisons among treatments for different traits

Treatment	Trait 1 (cm)	Trait 2 (%)	……	Trait n (g)
T1	12.1 ± 1.3 a [z]	5.0 ± 0.8 a	……	87 ± 5 a
T2	10.2 ± 1.1 a	4.9 ± 0.6 a	……	75 ± 3 ab
T3	3.9 ± 0.5 b	4.2 ± 0.5 a	……	67 ± 2 bc
T4	5.6 ± 1.0 b	10.9 ± 1.3 a	……	56 ± 3 c

[z] Mean ± standard error (n = 4). Means within each column followed by the same letter(s) are not significantly different at 5% level by Fisher's protected LSD test.

表 5. 經不同處理後採收果實之產量比較

Table 5. Comparisons of fruit yield among treatments

Treatment	Total yield (t/A)	Percentage of total yield (%) Marketable	Culls	Worm damaged
T1	17.0 ± 0.5 a [z]	85.3 ± 6.1 Aa	9.2 ± 0.6 Bc	5.5 ± 0.3 Cb
T2	16.1 ± 0.4 a	71.5 ± 4.3 Ab	23.2 ± 1.7 Ba	5.3 ± 0.5 Cb
T3	17.5 ± 0.7 a	70.7 ± 3.1 Ab	16.0 ± 2.1 Bb	13.3 ± 0.4 Ba
T4	17.1 ± 0.3 a	74.4 ± 5.8 Ab	12.5 ± 0.8 Bbc	13.1 ± 0.6 Ba

[z] Mean ± standard error (n = 4). Means within a column (in small letter) and within a row (in capital letter) followed by the same letter(s) are not significantly different at 5% level by Fisher's protected LSD test. Percentage data were arcsine-square-root transformed prior to analysis.

表 6. 氮肥施用量對不同品種之產量影響

Table 6. Effect of nitrogen treatment (N)[z] on yield of different varieties

	Yield (kg/ha)	
Variety (V)	N0	N120
V1	95.0 ± 4.1 aA [y]	94.2 ± 3.6 aA
V2	90.7 ± 5.3 aA	91.3 ± 3.7 bA
V3	86.5 ± 4.7 aA	89.5 ± 4.1 bA
V4	79.4 ± 3.8 bB	89.9 ± 2.8 bA
Source	F-test [x]	$LSD_{0.05}$ [w]
N	ns	–
V	*	5.6
N × V	**	4.8

[z] N0, without N treatment; N120, applying 120 kg/ha of N at flowering.
[y] Mean ± standard error (n = 3).　Means within a column (in small letter) and within a row (in capital letter) followed by the same letter(s) are not significantly different by Fisher's protected LSD test ($P < 0.05$).
[x] F-test of ANOVA. ns, non-significant; * and **, significant at 5% and 1% levels, respectively.
[w] $LSD_{0.05}$: values of least significant difference for mean comparisons at 5% level.

T2 和 T3 則差異未顯著。常見英文報告中以 *P* 值大小來闡釋顯著性結果，例如 It shows that the mean differences between T2 and T3 are not significant ($P < 0.05$)，根據英文文法結構，無論是不顯著或顯著，都應使用 $P < 0.05$ 而絕不可誤用 $P \leqq 0.05$ 或 $P > 0.05$；但在寫中文報告時，一般人會認為顯著或不顯著是以機率值大於 0.05 或小於 0.05 來判斷，從中文語法角度來看，若寫成：T2 和 T3 間差異未顯著 ($P < 0.05$)，似顯矛盾，於是直覺上就寫成：T2 和 T3 間差異未顯著 ($P > 0.05$)，但如此便造成同樣結論卻在中英文寫法上產生不一致的現象，而讓人混淆不清。因此，建議在撰寫中文報告時，最安全的作法是從顯著水準的角度來寫，例如不顯著之場合寫成：在 5%顯著水準下 T2 和 T3 間未達顯著差異，或 T2 和 T3 間差異未達 5%顯著水準；例如顯著之場合寫成：在 5%顯著水準下 T1 和 T2 之平均值皆顯著高於 T4，或 T1 和 T2 之平均值皆高於 T4 且其差異性達 5%顯著水準。

此外，由於執行 ANOVA 有其一定的統計前提，當不符合前提時，分析之前必須先行資料轉換 (data transformation) (Steel *et al*. 1997)。因此，若資料在分析前曾進行轉換，但統計表內所列數字是未經轉換的平均值時，務必在表下加註所採用的轉換方法，如表 5 所示，其中產量所佔比例之資料先經角度轉換後才進行 ANOVA 及 LSD 測驗，最後在表內以未經轉換之平均值來呈現。有關農學領域資料轉換的方法，可參閱 1990 年筆者翻譯 Finney (1989) 之論述報告而發表於國內「科學農業」期刊之文章 (Lu 1990b)。表 5 範例之分析結果可解讀為：經 LSD 測驗在 5%顯著水準下，四個處理對總產量並沒有顯著影響，但其中以 T1 處理能顯著提高可販售果實產量比例，剔出之次級品產量比例則以 T1 及 T4 處理最少而 T2 處理最多，T1 及 T2 兩處理能明顯降低蟲損產量比例，T3 及 T4 處理之次級品與蟲損產量比例間未有顯著差異等。

表 6 範例是將 ANOVA 及 LSD 測驗之分析結果同時呈現在同一表格內的常見用法之一，為求精簡，表中僅列出 ANOVA 之 F 測驗的顯著性，除顯著性符號仍使用星號*或**外，在此也將 F 測驗不顯著結果另以縮寫字母 ns 表示，以免留白時造成該處完全沒有任何數字或符號而讓人誤以為是漏字。從表 6 範例之分析結果，可知該兩因子試驗的氮肥及品種間交感作用 (N × V) 極顯著存在，表示氮肥對產量之影響會隨著不同品種而異，因此本例不能忽略氮肥效應而直接進行四個品種間之多重比較，或忽略品種效應而僅探討氮肥施用與否對所有品種產量之影響差異，而必須採取固定某因子變級對另一交感因子進行多重比較，即應分別就不施用及施用氮肥下各比較四個品種間之差異，反之也要就不同品種各自比較氮肥施用與

否間之差異，表 6 範例並將這兩種結果分別用逐欄及逐行比較呈現方式並列於表內。表 6 範例之分析結果可解讀為：從 ANOVA 之 F 測驗結果得知 N × V 交感作用極顯著存在，再經 LSD 測驗發現在 5%顯著水準下若不施氮肥，則 V1、V2 及 V3 三個品種之產量皆顯著高於 V4 品種，一旦施用每公頃 120 公斤氮肥，則 V1 品種之產量可顯著高於其他三個品種；有無施用氮肥對 V1、V2 及 V3 三品種之產量影響並不顯著，但施用氮肥可顯著提高 V4 品種之產量。

平均值比較表內也可增列 LSD 臨界值，如表 6 範例及表 7 實例 (Chan *et al.* 2008)，此時 LSD 的意義在表下仍應加註清楚；由於表 6 範例中的 N 效應不顯著，且 N × V 交感作用極顯著存在，直接比較 N 效應之差異性的意義不大，故其 LSD 臨界值可免列而以一個省略符號 (−) 來表示。表 7 實例是就各取樣時間下在不同處理內進行兩品種間之差異，故表下對於顯著性字母之意義的註解務必寫清楚，不能僅以上述 within each column 字眼即籠統帶過，否則會讓人誤以為是將欄內四個平均值一起進行比較。此處針對 mean ± standard error (n = 3) 的英文註解，也提供另一種寫法：Data indicate mean and its standard error of 3 replicates。

有時為了比較不同測量單位之各性狀間的變異程度，在處理平均值比較表內可列出變異係數 (coefficient of variation, CV) (Lu *et al.* 2005)，如表 8 所示，此時顯著性字母應該標於各平均值旁，而非 CV 值旁，這是因為無論何種性狀其 CV 的單位一律是%，與各性狀原測量單位並不相同 (而 SE 及 SD 與平均值具有相同的測量單位)，因此也不能用 ± 符號將平均值和 CV 值串在一起，即 CV 值必須單獨一欄。表下除顯著性字母及重複次數外，標題或表註也應清楚說明 CV 縮寫的定義。透過如表 8 範例之呈現方式，除了可知各性狀在四個處理間平均值及變異程度的差異外，亦可用來比較不同性狀間變異程度的差異，由表 8 顯示結果可解讀為：Trait n 的變異小於 Trait 1 和 Trait 2 兩性狀；因此各性狀 CV 值呈現的位數應一致，因而平均值位數在各性狀間亦維持一致，如表 8 範例所示，各性狀的平均值及 CV 值皆統一取至小數點後一位數字。有關 SE、SD 及 CV 的意義區分與使用場合，筆者曾於 2005 年

表 7. 有、無施行浸水處理下基因轉殖和非基因轉殖馬鈴薯收穫田區的殘留薯球數量比較 (摘自 Chan *et al.* 2008)

Table 7. Numbers of tubers of transgenic and non-transgenic plants of potatoes in harvested fields with or without flooding (cited from Chan *et al.* 2008)

		Number of tubers from 4 potato plants [z]				
Treatment	Cultivar [y]	7 [x]	17	28	51	70
Non- flooding	WT	14.3 ± 4.3 a	14.3 ± 2.1 a	14.3 ± 0.9 a	15.0 ± 1.5 a	14.7 ± 2.7 a
	2-1	10.3 ± 2.2 a	13.3 ± 1.2 a	8.7 ± 4.7 a	9.0 ± 3.2 a	5.7 ± 0.9 a
	LSD (0.05) [w]	14.9	12.5	16.9	15.1	12.4
Flooding	WT	6.0 ± 1.0 a	3.3 ± 2.8 a	0.0 ± 0.0 a	0.0 ± 0.0 a	0.0 ± 0.0 a
	2-1	6.0 ± 3.0 a	2.3 ± 1.2 a	0.0 ± 0.0 a	0.0 ± 0.0 a	0.0 ± 0.0 a
	LSD (0.05)	11.4	12.4	0.0	0.0	0.0

[z] Data indicate mean and its standard error of 3 replicates.
[y] WT: wild type (non-transgenic) potato (cv. Kennebec); 2-1: transgenic potato (strain 2-1).
[x] Sampling time: days of flooding treatment.
[w] LSD (0.05): values of least significant difference are for comparisons of cultivar means within treatments at 5% level. Within treatment for each sampling time, cultivar means followed by different letter are significantly different at 5% level by LSD test.

表 8. 多性狀在不同處理間平均值與變異係數之比較

Table 8. Comparison of mean and coefficient of variation (CV) among treatments for different traits

	Trait 1		Trait 2			Trait n	
Treatment	Mean (cm)	CV (%)	Mean (%)	CV (%)	……	Mean (g)	CV (%)
T1	12.1 a [z]	21.5	5.0 a	32.0	……	87.2 a	11.5
T2	10.2 a	21.6	4.9 a	24.5	……	75.0 ab	8.0
T3	3.9 b	25.6	4.2 a	23.8	……	67.1 bc	6.0
T4	5.6 b	35.7	10.9 a	23.9	……	55.8 c	10.7

[z] Data of each treatment are average of 4 replicates. Means within each column followed by the same letter(s) are not significantly different at 5% level by Fisher's protected LSD test.

在本所技術服務季刊有詳細說明 (Lu *et al.* 2005)。

多重比較的分析結果除了用表列方式外，也可改成使用圖示來表達，如長條圖 (如圖 1) 或折線圖 (如圖 2)，圖中可利用繪圖軟體或文書作業軟體內的插入文字功能，直接將經 LSD 測驗所得的顯著性字母標示其中，以同時呈現出處理間差異顯著性之結果。圖中的誤差線 (error bar)、重複次數及顯著性字母的意義也必須加註清楚，但須加註於圖標題內，而不像前述表格是加註於表下。由折線圖可同時看出性狀隨因子變化之趨勢，尤其適用於處理變級為數量值之場合，例如濃度變化。必須注意的是，由於 CV 的單位為%，異於各性狀原測量單位，因此圖中的誤差線，絕不可使用 CV 值，但可以 SE、SD 或平均值之信賴區間 (confidence interval, CI) 表示，端視結果解讀之目的來決定。長條圖的誤差線外觀可以同時標示上限、下限，如圖 1 所示，也可僅使用上限或下限，如圖 3 所示；但折線圖之誤差線通常同時標示上下限，如圖 2 所示。有關誤差線的選用原則及如何使用 Excel 製作，可參考筆者於 2009 年在本所技術服務季刊之發表報告 (Lu *et al.* 2009)。又，另以 LSD 測驗結果直接標示於統計圖中的其他範例，如圖 3 所示，由於圖 3 是用來解讀不同期作下三品系間之平均產量表現差異性，因此將 LSD 測驗所得之顯著性字母同時標於統計圖中，可清楚得知，經 LSD 測驗結果在 5%顯著水準下，春作以品系 L1 的平均產量顯著高於其他兩個品系，但在夏、秋兩作下三品系間的平均產量並未達顯著差異。

討　論

發表在學術期刊的研究報告，除研究成果之創新性及應用價值外，必須具備一定水準與品質，才有可能被接受刊登，如資料分析之正確性、圖表之品質以及結果解讀之精確度，皆是影響稿件是否能被接受的決定因素。表達資訊是統計圖表製作的首要原則，圖表的好與

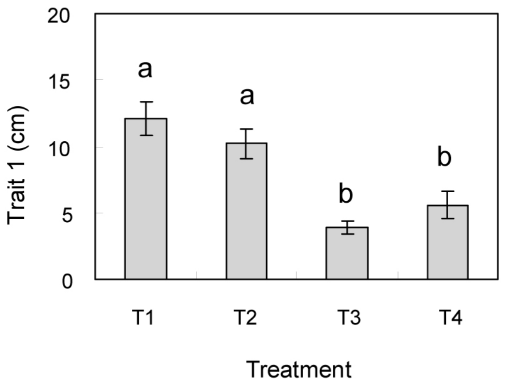

圖 1. 性狀 1 在處理間之比較。誤差線為平均值標準差 (n = 4)。各平均值上示以相同字母者為 5%水準下經 LSD 測驗未達顯著差異。(以長條圖為例)

Fig. 1. Comparison among treatments for Trait 1. Error bar is the standard error of mean (n = 4). Means with the same letter are not significantly different at 5% level by LSD test. (using a bar chart as an example)

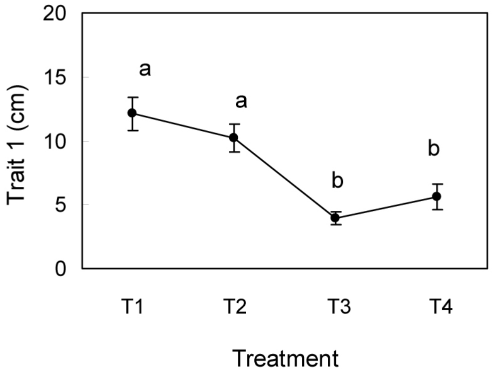

圖 2. 性狀 1 在處理間之比較。誤差線為平均值標準差 (n = 4)。各平均值上示以相同字母者為 5%水準下經 LSD 測驗未達顯著差異。(以折線圖為例)

Fig. 2. Comparison among treatments for Trait 1. Error bar is the standard error of mean (n = 4). Means with the same letter are not significantly different at 5% level by LSD test. (using a line graph as an example)

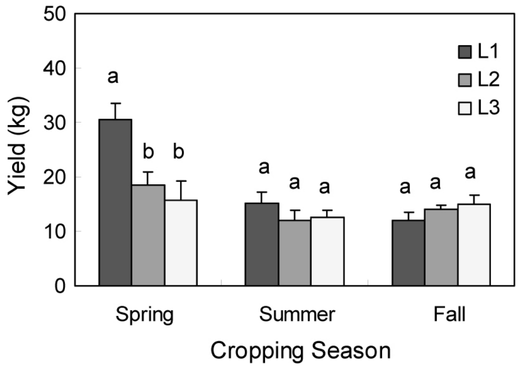

圖 3. 三個品系 (L1、L2 及 L3) 在不同期作下的產量比較。誤差線為平均值標準差 (n = 4)。各平均值上示以相同字母者為 5%水準下經 LSD 測驗未達顯著差異。

Fig. 3. Comparison among lines (L1, L2 and L3) for yields under different cropping seasons. Error bar is the standard error of mean (n = 4). Means with the same letter are not significantly different at 5% level by LSD test.

壞，足以左右作者對資料之判讀與讀者對報告品質之感觀。農業試驗資料最常見的統計分析為用以探明多組處理間差異的 ANOVA 及 LSD 測驗，ANOVA 是以一個表來呈現，而 LSD 可以用表或圖來表達。為了充分表達處理間比較之圖表所蘊含的意義，顯著性符號 (如 ANOVA 之星號及 LSD 之字母) 的正確使用及意義註解，甚為重要，此係研究人員在發表報告時最容易疏忽以致誤用之處。由於農業領域的試驗中往往會同時調查很多性狀，不同性狀就會有不同的 ANOVA 及 LSD 之結果呈現，如何在期刊有限篇幅內，將各性狀分析結果以更精簡有效的圖表來正確呈現並以精確文字來解讀，是每位研究人員應該具備的概念與基本能力。

引用文獻 (Literature cited)

Carmer, S. G. and W. M. Walker. 1985. Pairwise multiple comparisons of treatment means in agronomic research. J. Agron. Educ. 14:19–26.

Chan, L. F., M. T. Wu, and H. Y. Lu. 2008. Residue management at harvested fields of phytase containing transgenic potatoes: assessment of submerging treatment. J. Taiwan Agric. Res. 57: 95–102. (in Chinese with English abstract)

Chew, V. 1980. Testing differences among means: correct interpretation and some alternatives. HortScience 15:467–470.

Chiang, K. S. 2006. Analysis methods for multiple comparisons in plant protection research. Plant Prot. Bull. 48:259–268. (in Chinese with English abstract)

Finney, D. J. 1989. Was this in your statistics textbook? V. Transformation of data. Expl. Agric. 25:165–175.

Hsu, J. C. 1996. Multiple Comparisons. Chapman & Hall. New York. 277 pp.

Lu, H. Y. 1990a. Use and misuse of analysis methods for multiple comparisons. Sci. Agric. 38(7, 8):177–181. (in Chinese)

Lu, H. Y. 1990b. Was this in your statistics textbook? V. Transformation of data. Sci. Agric. 38 (7, 8): 182–187. (in Chinese)

Lu, H. Y., M. L. Wei, and C. T. Lu. 2005. Statistical analysis of agricultural research data using Excel. I. Descriptive statistics. Technol. Service of TARI 63:33–36. (in Chinese)

Lu, H. Y., M. L. Wei, and C. T. Lu. 2009. Statistical analysis of agricultural research data using Excel. VII. Error bar. Technol. Service of TARI 78:16–19. (in Chinese)

Steel, R. G. D., J. H. Torrie, and D. A. Dickey. 1997. Principles and Procedures of Statistics. 3rd ed. McGraw-Hill. New York, USA. 666 pp.

Streiner, D. L. 1996. Maintaining standards: differences between the standard deviation and standard error, and when to use each. Can. J. Psychiatry 41:498–502.

Correct Usage of Graph/Table for Treatment Comparisons[1]

Hsiu-Ying Lu[2,3]

Abstract

Lu, H. Y. 2011. Correct usage of graph/table for treatment comparisons. J. Taiwan Agric. Res. 60:61–71.

Conveying accurate information is key to the effective presentation of scientific graphs or tables. The analysis of variance (ANOVA) and multiple comparisons are the methods commonly used for comparison of several treatment means in data collection from agricultural research. There are many methods available for multiple comparisons of treatment means; among them, the Fisher's protected least significant difference (LSD) test is the favorable method for all pairwise comparisons of treatment means. A graph or table is of no scientific merit if the data are not analyzed statistically and presented accurately. The objective of this report is to describe proper procedure for preparation and presentation of statistically analyzed data in the form of scientific graphs or tables for improving quality of scientific publications.

Key words: Analysis of variance, Multiple comparisons, Pairwise comparisons of treatment means, Least significant difference test.

1. Contribution No. 2478 from Taiwan Agricultural Research Institute (TARI), Council of Agriculture. Accepted: March 23, 2011.
2. Researcher and Deputy Director General, TARI, Taichung, Taiwan, ROC.
3. Corresponding author, e-mail: iying@tari.gov.tw; Fax: (04)23302801.